

# NUMERICALLY ERASURE-ROBUST FRAMES

MATTHEW FICKUS AND DUSTIN G. MIXON

**ABSTRACT.** Given a channel with additive noise and adversarial erasures, the task is to design a frame that allows for stable signal reconstruction from transmitted frame coefficients. To meet these specifications, we introduce numerically erasure-robust frames. We first consider a variety of constructions, including random frames, equiangular tight frames and group frames. Later, we show that arbitrarily large erasure rates necessarily induce numerical instability in signal reconstruction. We conclude with a few observations, including some implications for maximal equiangular tight frames and sparse frames.

## 1. INTRODUCTION

Modern communication networks are rooted in both information theory and algebraic coding theory. In these contexts, after deciding on a probabilistic noise model for a given communication channel, one chooses an appropriate error-correcting code to achieve reliable communication with a maximal information rate. For linear codes in particular, encoding and decoding reduce to problems in linear algebra over finite fields. Beginning with Goyal et al. [16], finite frame theorists have studied the generalizations of these problems to real and complex vector spaces. This generalization allows one to use certain mathematical tools, such as matrix norms and condition numbers, which are not well-defined in the finite-field setting.

This paper is concerned with a channel characterized by additive noise and adversarial erasures. We encode a signal  $x \in \mathbb{C}^M$  using inner products  $\langle x, f_n \rangle$  with members of a spanning sequence of vectors  $\{f_n\}_{n=1}^N \subseteq \mathbb{C}^M$ ; such a sequence is called a *frame*. In transmitting these inner products, we expect additive noise due to various phenomena such as atmospheric interactions or round-off error. If these were the only sources of noise, then it would be reasonable to reconstruct the original signal by applying the Moore-Penrose pseudoinverse. To be precise, letting  $F$  denote the  $M \times N$  matrix whose columns are the  $f_n$ 's, we transmit  $y = F^*x$ . At the receiver, an estimate of  $x$  is then found by computing

$$\hat{x} = ((FF^*)^{-1}F)(y + e) = x + (FF^*)^{-1}Fe,$$

where  $e$  is additive noise. Assuming the channel has a “signal-to-noise ratio” of  $R = \|y\|/\|e\|$ , we can estimate how the size of the estimate error  $(FF^*)^{-1}Fe$  compares with the size of the

---

2000 *Mathematics Subject Classification.* 42C15, 15A12.

*Key words and phrases.* frames, erasures, well-conditioned.

The authors thank the anonymous referee for very helpful comments and suggestions. MF was supported by NSF Grant No. DMS-1042701 and AFOSR Grant Nos. F1ATA01103J001 and F1ATA00183G003, and DGM was supported by the A.B. Krongard Fellowship. The views expressed in this article are those of the authors and do not reflect the official policy or position of the United States Air Force, Department of Defense, or the U.S. Government.

original signal  $x$ . Indeed,  $\|(FF^*)^{-1}Fe\| \leq \frac{C}{R}\|x\|$ , where

$$C := \sup_{\substack{x \in \mathbb{C}^M \setminus \{0\} \\ e \in \mathbb{C}^N \setminus \{0\}}} R \cdot \frac{\|(FF^*)^{-1}Fe\|}{\|x\|} = \sup_{\substack{x \in \mathbb{C}^M \setminus \{0\} \\ e \in \mathbb{C}^N \setminus \{0\}}} \frac{\|F^*x\|}{\|x\|} \cdot \frac{\|(FF^*)^{-1}Fe\|}{\|e\|} = \|F\|_2 \|(FF^*)^{-1}F\|_2.$$

Here,  $C$  is the *condition number* of  $F$ , denoted  $\text{Cond}(F)$ , which is equal to the ratio of the greatest singular value of  $F$  to its smallest one. From this perspective, the best possible frames are those with  $\text{Cond}(F) = 1$ , a fact which occurs precisely when  $FF^* = AI_M$  for some  $A > 0$ ; such  $F$ 's are called *tight frames*.

We consider channels that, in addition to additive noise, suffer from *erasures*. To be precise, the transmitted signal is a sequence of inner products:  $F^*x = \{\langle x, f_n \rangle\}_{n=1}^N$ . Like [16], we consider channels which completely delete some of these inner products and add noise to the remaining ones. However, whereas [16] focuses on average reconstruction performance, we instead follow [11] and [17] by focusing on worst-case reconstruction performance. In particular, by considering worst-case performance, we design frames which are robust against the erasure of *any* fixed number of inner products. Such frames could be particularly useful in situations where an adversary is actively deleting our most useful frame coefficients, i.e., active jamming. We say that such frames are robust against *adversarial* erasures.

To design such frames, we first acknowledge that we cannot reconstruct the  $M$ -dimensional signal  $x$  without at least  $M$  inner products. As such, we must impose some constraint on the adversary. For the highly constrained adversary, Casazza and Kovačević [11] show that tight frames of unit-norm vectors, called *unit norm tight frames*, are optimally robust against one erasure. Soon thereafter, Holmes and Paulsen [17] showed that *equiangular tight frames*—explicitly defined in the next section—are optimal for two erasures. To combat the highly destructive adversary, Püschel and Kovačević [20] propose frames which are *maximally robust* to erasures in the sense that the original signal can be recovered from any  $M$  of the  $N$  inner products. Other constructions of such *maximally robust frames* are given in [1], where they are dubbed *full spark frames*. It remains unclear whether the deletion of any  $N - M$  frame coefficients will allow for numerically stable reconstruction; this is an important distinction between invertible submatrices—the subject of [1, 20]—and well-conditioned submatrices, which is our focus here.

To be clear, in this paper we consider the case where the adversary is only capable of removing a proportion  $p$  of the  $N$  transmitted inner products. Then the remaining  $(1 - p)N$  inner products correspond to a subcollection of  $(1 - p)N$  columns of  $F$ , which we require to be well-conditioned for our reconstruction to properly combat the additive noise. Since erasures occur according to the will of an adversary, as opposed to a random process, we must ensure that *every* subcollection of  $(1 - p)N$  columns of  $F$  is well-conditioned. This leads to the following definition:

**Definition 1.** Given  $p \in [0, 1]$  and  $C \geq 1$ , an  $M \times N$  frame  $F$  is a  $(p, C)$ -numerically erasure-robust frame (NERF) if for every  $\mathcal{K} \subseteq \{1, \dots, N\}$  of size  $K := (1 - p)N$ , the corresponding  $M \times K$  submatrix  $F_{\mathcal{K}}$  has condition number  $\text{Cond}(F_{\mathcal{K}}) \leq C$ .

The purpose of this paper is to make the first strides in studying NERFs. In the following section, we use a variety of techniques to form different NERF constructions. Taking

inspiration from matrix design problems in compressed sensing, we first investigate frames whose entries are independent Gaussian random variables. Next, we consider equiangular tight frames, with which we get stronger results at the price of higher redundancy in the frame. Later, we show how the symmetry of group frames makes them naturally amenable to NERF analysis. In Section 3, we report a result on the fundamental limits of NERFs: that NERFs cannot stably support erasure rates  $p$  which are arbitrarily close to 1. Finally, we conclude with a few interesting observations in Section 4.

## 2. CONSTRUCTIONS

**2.1. Random frames.** The reader may have noticed some similarity between the definition of numerically erasure-robust frames and a matrix property which comes from the compressed sensing literature: the *restricted isometry property (RIP)*. To be clear, an  $M \times N$  matrix  $F$  is RIP if it acts as a near-isometry on sufficiently sparse vectors, that is,  $\|Fx\| \approx \|x\|$  for all vectors  $x$  with sufficiently few nonzero entries [12]. In other words, submatrices  $F_{\mathcal{K}}$  composed of sufficiently few columns from  $F$  have  $F_{\mathcal{K}}^* F_{\mathcal{K}}$  particularly close to the identity matrix, meaning  $F_{\mathcal{K}}^* F_{\mathcal{K}}$  is particularly well-conditioned. The key difference between NERFs and RIP matrices is that well-conditioned NERF submatrices  $F_{\mathcal{K}}$  have  $K := |\mathcal{K}| \geq M$  columns, whereas  $F_{\mathcal{K}}$  has fewer than  $M$  columns in the RIP case. Regardless, in constructing NERFs, we can exploit some intuition from the construction of RIP matrices. In particular, the RIP matrices which support the largest sparsity levels to date arise from random processes. As an example, one may draw the entries independently from a Gaussian distribution of mean zero and variance  $\frac{1}{M}$ ; this was originally established in Lemma 3.1 of [13]. What follows is the analogous NERF result:

**Theorem 2.** *Fix  $\varepsilon > 0$  and pick an  $M \times N$  frame  $F$  by drawing each entry independently from a standard normal distribution. Then  $F$  is a  $(p, C)$ -numerically erasure-robust frame with overwhelming probability provided*

$$\sqrt{\frac{M}{N}} \leq \frac{C-1}{C+1} \sqrt{1-p} - \sqrt{\varepsilon + 2p(1-\log p)}. \quad (1)$$

Note that (1) requires its right-hand side to be positive, which in turn implies

$$\sqrt{1-p} - \sqrt{2p(1-\log p)} > 0.$$

This occurs whenever  $p \leq 0.1460$ . That is, the random construction in Theorem 2 is numerically robust to erasure rates of up to almost 15%. However, approaching a 15% erasure rate while satisfying (1) will admittedly cost a large worst-case condition number  $C$  along with high redundancy  $\frac{N}{M}$  in the frame. Still, Theorem 2 provides a useful guarantee. For example, a Gaussian matrix of redundancy  $\frac{N}{M} = 5$  will, with overwhelming probability, be robust to 1% erasures with a worst-case condition number of 10. We proceed with the proof:

*Proof of Theorem 2.* Pick  $\mathcal{K} \subseteq \{1, \dots, N\}$  of size  $K = (1-p)N$ . Note the assumption (1) implies that  $\frac{M}{N} \leq 1-p$  and so  $K = (1-p)N \geq M$ . As such, Theorem II.13 of [14] gives bounds on the singular values of the random “tall”  $K \times M$  matrix  $F_{\mathcal{K}}^*$ :

$$\Pr[\sqrt{K} - \sqrt{M} - t \leq \sigma_{\min}(F_{\mathcal{K}}^*) \leq \sigma_{\max}(F_{\mathcal{K}}^*) \leq \sqrt{K} + \sqrt{M} + t] \geq 1 - 2e^{-t^2/2} \quad \forall t \geq 0.$$

This probabilistic bound on the extreme singular values implies

$$\Pr \left[ \text{Cond}(F_{\mathcal{K}}) \leq \frac{\sqrt{K} + \sqrt{M} + t}{\sqrt{K} - \sqrt{M} - t} \right] \geq 1 - 2e^{-t^2/2} \quad \forall t \geq 0.$$

Taking a union bound over all  $\binom{N}{K} = \binom{N}{N-K} \leq (\frac{eN}{N-K})^{N-K}$  choices for  $\mathcal{K}$  gives

$$\begin{aligned} \Pr \left[ \exists \mathcal{K} \text{ s.t. } \text{Cond}(F_{\mathcal{K}}) > \frac{\sqrt{K} + \sqrt{M} + t}{\sqrt{K} - \sqrt{M} - t} \right] &\leq \binom{N}{N-K} 2e^{-t^2/2} \\ &\leq 2 \exp \left( -\frac{t^2}{2} + (N-K) \log \frac{eN}{N-K} \right) \\ &= 2 \exp \left( -\frac{t^2}{2} + Np \log \frac{e}{p} \right) \quad \forall t \geq 0. \end{aligned} \quad (2)$$

Now pick  $t$  such that  $C = \frac{\sqrt{K} + \sqrt{M} + t}{\sqrt{K} - \sqrt{M} - t}$ , namely  $t = \sqrt{N}(\frac{C-1}{C+1}\sqrt{1-p} - \sqrt{\frac{M}{N}})$ . Note that (1) implies  $t \geq 0$ , and so we may substitute it into (2) and simplify the result:

$$\begin{aligned} \Pr [\exists \mathcal{K} \text{ s.t. } \text{Cond}(F_{\mathcal{K}}) > C] &\leq 2 \exp \left[ -\frac{N}{2} \left( \left( \frac{C-1}{C+1} \sqrt{1-p} - \sqrt{\frac{M}{N}} \right)^2 - 2p(1 - \log p) \right) \right] \\ &\leq 2e^{-N\varepsilon/2}. \end{aligned}$$

Thus, the probability of  $F$  not being a  $(p, C)$ -NERF is  $O(N^{-\alpha})$  for every fixed  $\alpha$ , meaning  $F$  is a  $(p, C)$ -NERF with overwhelming probability.  $\square$

**2.2. Equiangular tight frames.** The previous subsection constructed a random family of numerically erasure-robust frames by following intuition from known constructions of matrices with the restricted isometry property. Indeed, state-of-the-art RIP matrices are built according to random processes, while deterministic constructions have found less success [5]. In this subsection, the analogy between RIP matrices and NERFs will break down, as we will construct deterministic NERFs which outperform the random counterparts with much larger erasure rates, albeit at the price of high redundancy.

In [17], Holmes and Paulsen show that frames of pairwise dissimilar unit-norm vectors are robust to two erasures. This dissimilarity is measured in terms of worst-case coherence, which is known to satisfy the Welch bound:

**Theorem 3** (Welch bound [25]). *Every  $M \times N$  frame  $\{f_n\}_{n=1}^N$  of unit-norm vectors has worst-case coherence*

$$\max_{\substack{n, n' \in \{1, \dots, N\} \\ n \neq n'}} |\langle f_n, f_{n'} \rangle| \geq \sqrt{\frac{N-M}{M(N-1)}}.$$

Specifically, Proposition 2.2 of [17] gives that minimizers of worst-case coherence are optimally robust to two erasures. For many values of  $M$  and  $N$ , there exist frames which achieve equality in the Welch bound. In fact, a sequence of unit-norm vectors  $F = \{f_n\}_{n=1}^N$  achieves the Welch bound if and only if it is an *equiangular tight frame (ETF)*, meaning that it is a tight frame (i.e.,  $FF^* = AI_M$ ) which also satisfies the equiangularity condition

that  $|\langle f_n, f_{n'} \rangle|$  is constant over all choices of  $n \neq n'$  [22]. Not only are ETFs minimizers of worst-case coherence, they also have combinatorial symmetries related to strongly regular graphs, difference sets and Steiner systems; these combinatorial structures have each been used to build the only general ETF constructions to date [15, 24, 26].

In this subsection, we consider an ETF construction based on a particular difference set. Let  $q$  be a prime power, take  $M = q + 1$  and  $N = q^2 + q + 1$ , and consider the trace map  $\text{Tr} : \mathbb{F}_{q^3} \rightarrow \mathbb{F}_q$  defined by  $\text{Tr}(\beta) = \beta + \beta^q + \beta^{q^2}$ . Given a generator  $\alpha$  of the multiplicative group of  $\mathbb{F}_{q^3}$ , define the  $M$ -element subset  $\mathcal{M} \subseteq \mathbb{Z}_N$  by  $\mathcal{M} = \{t : \text{Tr}(\alpha^t) = 0\}$ . By construction,  $\mathcal{M}$  has the property that every nonzero member of  $\mathbb{Z}_N$  can be uniquely expressed as the difference of two elements of  $\mathcal{M}$ ; this set is called the  $(N, M, 1)$ -Singer difference set [18]. As shown in [26], any difference set  $\mathcal{M} \subseteq \mathbb{Z}_N$  can be used to build an ETF by taking rows from the  $N \times N$  discrete Fourier transform matrix which are indexed by members of  $\mathcal{M}$  and then normalizing the resulting columns. This construction has the following guarantee:

**Theorem 4.** *Take  $M = q + 1$  and  $N = q^2 + q + 1$  for some prime power  $q$ , and let  $F$  be the  $M \times N$  equiangular tight frame  $F$  constructed from the  $(N, M, 1)$ -Singer difference set, as in [26]. Then  $F$  is a  $(p, C)$ -numerically erasure-robust frame for every  $p \leq \frac{1}{2} - \frac{C^2}{C^4 + 1}$ .*

This result essentially states that such ETFs are numerically robust to erasure rates of up to 50%. Compared to the random construction of the previous section, which required less than 15% erasures, this is quite an improvement. Certainly, the frame redundancy  $\frac{N}{M}$  is unbounded in this case since  $N$  scales as  $M^2$ , but the reward is significant. For example, such ETFs are robust to 49% erasures with a worst-case condition number of 10. Meanwhile, for  $N \gg M$ , Theorem 2 only guarantees—with overwhelming probability—a worst-case condition number of 10 when less than 9% of the frame is erased.

*Proof of Theorem 4.* Pick some  $\mathcal{K} \subseteq \{1, \dots, N\}$  of size  $K = (1 - p)N$ , and let  $\{\lambda_{\mathcal{K};m}\}_{m=1}^M$  denote the eigenvalues of  $F_{\mathcal{K}}F_{\mathcal{K}}^*$ . Taking  $\delta_{\mathcal{K}} := \max_m |\frac{M}{K}\lambda_{\mathcal{K};m} - 1|$ , we have

$$(\text{Cond}(F_{\mathcal{K}}))^2 = \text{Cond}(F_{\mathcal{K}}F_{\mathcal{K}}^*) = \frac{\lambda_{\max}(F_{\mathcal{K}}F_{\mathcal{K}}^*)}{\lambda_{\min}(F_{\mathcal{K}}F_{\mathcal{K}}^*)} \leq \frac{1 + \delta_{\mathcal{K}}}{1 - \delta_{\mathcal{K}}} \quad (3)$$

provided  $\delta_{\mathcal{K}} < 1$ ; if  $\delta_{\mathcal{K}} \geq 1$ , then  $F_{\mathcal{K}}$  could be rank deficient. Moreover, the fact that  $F_{\mathcal{K}}F_{\mathcal{K}}^*$  and  $\mathbf{I}_M$  are simultaneously diagonalizable implies

$$\delta_{\mathcal{K}}^2 = \frac{M^2}{K^2} \max_{m \in \{1, \dots, M\}} |\lambda_{\mathcal{K};m} - \frac{K}{M}|^2 \leq \frac{M^2}{K^2} \sum_{m=1}^M |\lambda_{\mathcal{K};m} - \frac{K}{M}|^2 = \frac{M^2}{K^2} \text{Tr}[(F_{\mathcal{K}}F_{\mathcal{K}}^* - \frac{K}{M}\mathbf{I}_M)^2]. \quad (4)$$

From here, the cyclic property of the trace and the fact that  $F$  has unit-norm columns give

$$\begin{aligned} \text{Tr}[(F_{\mathcal{K}}F_{\mathcal{K}}^* - \frac{K}{M}\mathbf{I}_M)^2] &= \text{Tr}[(F_{\mathcal{K}}F_{\mathcal{K}}^*)^2] - \frac{2K}{M}\text{Tr}[F_{\mathcal{K}}F_{\mathcal{K}}^*] + \frac{K^2}{M^2}\text{Tr}[\mathbf{I}_M] \\ &= \text{Tr}[(F_{\mathcal{K}}^*F_{\mathcal{K}})^2] - \frac{2K}{M}\text{Tr}[F_{\mathcal{K}}^*F_{\mathcal{K}}] + \frac{K^2}{M} \\ &= \sum_{k \in \mathcal{K}} \sum_{k' \in \mathcal{K}} |\langle f_k, f_{k'} \rangle|^2 - \frac{K^2}{M}. \end{aligned} \quad (5)$$

Since  $F$  is an ETF, the inner products between distinct frame elements achieve equality in the Welch bound:  $|\langle f_k, f_{k'} \rangle|^2 = \frac{N-M}{M(N-1)}$  for every  $k \neq k'$ . Applying this to (5) and substituting

into (4) then gives

$$\delta_{\mathcal{K}}^2 \leq \frac{M^2}{K^2} \left( K + K(K-1) \frac{N-M}{M(N-1)} - \frac{K^2}{M} \right) = \frac{M(M-1)(N-K)}{K(N-1)} = \frac{pM(M-1)}{(1-p)(N-1)}. \quad (6)$$

According to the theorem statement,  $N = M^2 - M + 1$  and  $p \leq \frac{1}{2} - \frac{C^2}{C^4+1}$ , and so

$$\delta_{\mathcal{K}}^2 \leq \frac{p}{1-p} \leq \frac{(C^2-1)^2}{(C^2+1)^2}.$$

Substituting this into (3) therefore gives  $\text{Cond}(F_{\mathcal{K}}) \leq C$ .  $\square$

We note that (6) together with the necessary condition  $\delta_{\mathcal{K}}^2 < 1$  indicate that of all  $M \times N$  ETFs, the above proof technique will only work for those with  $N = \Omega(M^2)$  frame elements. However, as noted in Proposition 2.3 of [4],  $M \times N$  ETFs necessarily have  $N \leq M^2$ , and so the ETFs for which the above proof can demonstrate NERF are *asymptotically maximal*. A long-standing open problem in frame theory concerns the existence of  $M \times N$  ETFs with  $N = M^2$ , or *maximal ETFs*, and it is easy to verify that Theorem 4 also holds for this conjectured family; to date, these are only known to exist for finitely many  $M$ 's [3]. As for asymptotically maximal ETFs, the difference set construction of Theorem 4 is the only such infinite family known to the authors. Regardless, a version of Theorem 4 holds for every family of asymptotically maximal ETFs, which follows directly from (6):

**Theorem 5.** *Every  $M \times N$  equiangular tight frame with  $\frac{N-1}{M(M-1)} \geq \alpha$  is a  $(p, C)$ -numerically erasure-robust frame for every  $p \leq \frac{\alpha(C^2-1)^2}{\alpha(C^2-1)^2 + (C^2+1)^2}$ .*

Since maximal ETFs are particularly difficult to construct, different fields have turned to *mutually unbiased bases (MUBs)* to fill their need for large frames with low coherence [19, 22]. There are several  $M \times M^2$  MUB constructions, all of which have the property that the inner product between any two columns is of size 0 or  $1/\sqrt{M}$  [2, 9, 19]. As the Welch bound in this case is  $1/\sqrt{M+1}$ , MUBs are “almost” ETFs. It is therefore surprising that the above proof techniques fail to show that MUBs are NERFs. To illustrate this fact, we consider the MUB version of (6):

$$\delta_{\mathcal{K}}^2 \leq \frac{M^2}{K^2} \left( K + K(K-1) \frac{1}{M} - \frac{K^2}{M} \right) = \frac{M(M-1)}{K} = \frac{M-1}{(1-p)M}. \quad (7)$$

Due to the necessity of  $\delta_{\mathcal{K}} < 1$ , this bound will not be useful unless  $p < \frac{1}{M}$ . However, even in this case, substituting (7) into (3) gives

$$(\text{Cond}(F_{\mathcal{K}}))^2 \leq \frac{1 + \delta_{\mathcal{K}}}{1 - \delta_{\mathcal{K}}} \leq \frac{\sqrt{(1-p)M} + \sqrt{M-1}}{\sqrt{(1-p)M} - \sqrt{M-1}}. \quad (8)$$

Further since  $0 \leq p \leq \frac{1}{M}$ , separately bounding the numerator and denominator gives that the right-hand side of (8) is always at least  $2\sqrt{M-1}$ , meaning (8) says very little about the worst-case condition number, regardless of the erasure rate.

It remains to be seen whether this is a true distinction between ETFs and MUBs or is instead an artifact of our proof techniques. One way to improve this analysis is to find a better bound on the *frame potential* (5), see [7]. To be clear, we can certainly bound it in general

using worst-case coherence, and such a bound is tight whenever the frame is equiangular. However, when the frame is not equiangular, this bound is less than optimal. For a better bound in the general case, suppose that for every  $n \in \{1, \dots, N\}$ , the distribution of the squares of inner products  $\{|\langle f_n, f_{n'} \rangle|^2\}_{n'=1}^N$  is identical. In this case, let  $d_F \in \mathbb{R}^N$  denote the common sequence of squared inner products, sorted in nonincreasing order. We can then bound the sum in (5) by exploiting this structure:

$$\sum_{k \in \mathcal{K}} \sum_{k' \in \mathcal{K}} |\langle f_k, f_{k'} \rangle|^2 \leq K \sum_{k=1}^K d_F[k]. \quad (9)$$

Combining bounds (4), (5) and (9) then yields

$$\delta_{\mathcal{K}}^2 \leq \frac{M^2}{K^2} \left( \sum_{k \in \mathcal{K}} \sum_{k' \in \mathcal{K}} |\langle f_k, f_{k'} \rangle|^2 - \frac{K^2}{M} \right) \leq M^2 \left( \frac{1}{K} \sum_{k=1}^K d_F[k] - \frac{1}{M} \right). \quad (10)$$

In particular, in order to use (10) to guarantee  $\delta_{\mathcal{K}} < 1$ , we want the average of the  $K$  largest values of  $d_F[k]$  to be close to  $\frac{1}{M}$ . Further note that if  $F$  is a unit norm tight frame, which necessarily has tight frame constant  $A = \frac{N}{M}$ , then the average of all values of  $d_F[k]$  is  $\frac{1}{M}$ :

$$\frac{1}{N} \sum_{k=1}^N d_F[k] = \frac{1}{N} \sum_{n=1}^N |\langle f_n, f_{n'} \rangle|^2 = \frac{1}{N} \frac{N}{M} \|f_{n'}\|^2 = \frac{1}{M}.$$

In such cases, using (10) to estimate the NERF properties of a given frame reduces to finding how quickly (as a function of  $K$ ) the average of the  $K$  largest values of  $d_F[k]$  converges to the average of all of its values.

With this refined analysis, we can prove that MUBs are actually NERFs. We note that the bound (9) is identical to the worst-case coherence bound unless  $K$  is large, since  $d_F$  in this case has one copy of 1,  $M(M-1)$  copies of  $\frac{1}{M}$ , and  $M-1$  copies of 0 [2, 9, 19]. Indeed, analysis with (9) can only show that MUBs are NERFs when the erasure rate is small:

**Theorem 6.** *An  $M \times M^2$  frame of mutually unbiased bases is a  $(p, C)$ -numerically erasure-robust frame for every  $p \leq \frac{(C^2-1)^2}{(C^2+1)^2(M+1)}$ .*

Note that the above guarantee is not nearly as good as the one we got for ETFs, or even for random frames. However, the result is still of some use; for example, when  $M$  is sufficiently large, removing any  $0.96M$  of the  $M^2$  frame vectors will leave a submatrix of condition number smaller than 10.

*Proof of Theorem 6.* Applying (10) to the distribution  $d_F$  of the  $M \times M^2$  MUB yields

$$\delta_{\mathcal{K}}^2 \leq \frac{M^2}{K} \left( \sum_{k=1}^K d_F[k] - \frac{K}{M} \right) = \frac{M^2}{K} \left( 1 + M(M-1) \frac{1}{M} - \frac{K}{M} \right) = \frac{M(M^2 - K)}{K}.$$

Since  $K = (1 - p)N$  and  $N = M^2$ , we can simplify and apply  $p \leq \frac{(C^2-1)^2}{(C^2+1)^2(M+1)}$  to get

$$\begin{aligned} \delta_K^2 &\leq \frac{pM}{1-p} \leq \frac{(C^2-1)^2 M}{(C^2+1)^2(M+1) - (C^2-1)^2} \\ &\leq \frac{(C^2-1)^2 M}{(C^2+1)^2(M+1) - (C^2+1)^2} = \frac{(C^2-1)^2}{(C^2+1)^2}. \end{aligned} \quad (11)$$

Substituting this into (3) therefore gives  $\text{Cond}(F_K) \leq C$ .  $\square$

**2.3. Group frames.** In the previous subsection, we demonstrated that mutually unbiased bases are NERFs by exploiting an important property: the distribution of the squares of inner products  $\{|\langle f_n, f_{n'} \rangle|^2\}_{n'=1}^N$  is identical for every  $f_n$ . In this subsection, we will consider a much larger class of unit norm tight frames that enjoy this identical distribution property: group frames. Given a seed vector  $f \in \mathbb{C}^M$  and a finite subgroup  $G$  of the group of all  $M \times M$  unitary matrices, the corresponding *group frame* is the orbit  $\{Uf\}_{U \in G}$  of  $f$  under the action of this group, though  $\{Uf\}_{U \in G}$  should only be called a frame if the  $Uf$ 's span. In fact, if  $\|f\| = 1$ , then  $\{Uf\}_{U \in G}$  will be a unit norm tight frame provided the group  $G$  is *irreducible*, meaning that for any nonzero  $x \in \mathbb{C}^M$  the vectors  $\{Ux\}_{U \in G}$  necessarily span  $\mathbb{C}^M$ ; for this and other interesting facts about group frames, see [23]. Note that for any  $U, U' \in G$ ,

$$\langle Uf, U'f \rangle = \langle f, U^*U'f \rangle = \langle f, U^{-1}U'f \rangle.$$

Since each  $U^{-1}$  acts as a permutation on  $G$ , we conclude that  $\{\langle Uf, U'f \rangle\}_{U' \in G}$  is a permutation of  $\{\langle f, U'f \rangle\}_{U' \in G}$ , thereby confirming our above claim that each row of the Gram matrix  $F^*F$  is identically distributed.

To illustrate the usefulness of group frame ideas in estimating  $\delta_K$  with (10), we will apply it to group frames generated by the symmetric group of the simplex. First, we define a (*regular*) *simplex* to be any  $M \times (M+1)$  matrix  $\Psi$  whose  $(M+1) \times (M+1)$  Gram matrix is  $\Psi^*\Psi = \frac{M+1}{M}I_{M+1} - \frac{1}{M}J_{M+1}$ , where  $J_{M+1}$  denotes an  $(M+1) \times (M+1)$  matrix of ones. Notice that the spectrum of  $\Psi^*\Psi$  consists of  $M$  copies of  $\frac{M+1}{M}$  and one value of 0; since this is a zero-padded version of the spectrum of the  $M \times M$  frame operator  $\Psi\Psi^*$ , we conclude that  $\Psi\Psi^* = \frac{M+1}{M}I_M$ , meaning  $\Psi$  is a tight frame. In fact, since the off-diagonal entries of  $\Psi^*\Psi$  are all equal in size (to the Welch bound),  $\Psi$  is an equiangular tight frame.

The simplex plays an important role in finite frame theory. Indeed, the Mercedes-Benz frame and the vertices of the tetrahedron, being 2- and 3-dimensional realizations of the simplex, serve as fundamental examples of frames [7, 23]. Simplices can also be easily expressed in higher dimensions by removing the row of 1's from an  $(M+1) \times (M+1)$  discrete Fourier transform matrix or Hadamard matrix and then normalizing the resulting columns. This representation of simplices plays a key role in the construction of Steiner ETFs [15]. In this paper, we are specifically interested in the *symmetries* of the simplex. In general, the *symmetry group* of a frame is the set of all matrices which, when acting on frame elements, permute them. The following result gives a particularly nice description of the symmetry group of the simplex:

**Lemma 7.** *The symmetry group of an  $M \times (M+1)$  regular simplex  $\Psi$  is the set of all matrices of the form  $U = \frac{M}{M+1}\Psi P\Psi^*$ , where  $P$  is an  $(M+1) \times (M+1)$  permutation matrix.*



*Proof.* The symmetry group of  $\Psi$  is the set of all matrices  $U$  for which there exists a permutation matrix  $P$  such that  $U\Psi = \Psi P$ . Note this implies  $U\Psi\Psi^* = \Psi P\Psi^*$  which, since  $\Psi\Psi^* = \frac{M+1}{M}\mathbf{I}_M$ , further implies  $U = \frac{M}{M+1}\Psi P\Psi^*$ . In other words, for each member  $U$  of the symmetry group of  $\Psi$ , there is a unique permutation matrix  $P$  such that  $U\Psi = \Psi P$ . Thus, all that remains to be shown is that for each permutation matrix  $P$ , the matrix  $U = \frac{M}{M+1}\Psi P\Psi^*$  satisfies  $U\Psi = \Psi P$ . To this end, note

$$U\Psi = \frac{M}{M+1}\Psi P\Psi^*\Psi = \frac{M}{M+1}\Psi P\left(\frac{M+1}{M}\mathbf{I}_{M+1} - \frac{1}{M}\mathbf{J}_{M+1}\right) = \Psi P - \frac{1}{M+1}\Psi P\mathbf{J}_{M+1}.$$

It therefore suffices to show that  $\Psi P\mathbf{J}_{M+1} = 0$ . To do this, factor  $\mathbf{J}_{M+1}$  as an outer product of an all-ones vector with itself, a vector which happens to be preserved by permutations:  $\Psi P\mathbf{J}_{M+1} = \Psi P\mathbf{1}_{M+1}\mathbf{1}_{M+1}^* = \Psi\mathbf{1}_{M+1}\mathbf{1}_{M+1}^*$ . Then note that  $\Psi\mathbf{1}_{M+1} = 0$ :

$$\|\Psi\mathbf{1}_{M+1}\|^2 = \mathbf{1}_{M+1}^*\Psi^*\Psi\mathbf{1}_{M+1} = \mathbf{1}_{M+1}^*\left(\frac{M+1}{M}\mathbf{I}_{M+1} - \frac{1}{M}\mathbf{1}_{M+1}\mathbf{1}_{M+1}^*\right)\mathbf{1}_{M+1} = 0. \quad \square$$

From Lemma 7, we can deduce that the symmetry group of an  $M \times (M+1)$  simplex  $\Psi$  is the symmetric group on  $M+1$  letters, and so we denote it by  $S_{M+1}$ . We are interested in the frames formed by applying the  $(M+1)!$  members of  $S_{M+1}$  to unit vectors. We claim that such frames are automatically unit norm tight frames. Moreover, motivated by (10), we further seek the distribution  $d_F$  of the squared-moduli of the inner products of the frame elements with each other.

Here, it is helpful to note that  $\Phi^* := \sqrt{M/(M+1)}\Psi^*$  is a unitary transformation between  $\mathbb{C}^M$  and the  $M$ -dimensional orthogonal complement  $1_{M+1}^\perp$  of the  $(M+1)$ -dimensional all-ones vector; the proof of this fact is straightforward and is not included here. Indeed, writing any unit-norm vector  $f \in \mathbb{C}^M$  as  $f = \Phi g$  where  $g \in 1_{M+1}^\perp$  has  $\|g\| = 1$ , we have inner products of the form:

$$\langle f, Uf \rangle = \langle f, \frac{M}{M+1}\Psi P\Psi^* f \rangle = \langle \Phi^* f, P\Phi^* f \rangle = \langle g, Pg \rangle. \quad (12)$$

Moreover, as noted above, our group frame will be tight provided that for any  $x \neq 0$  the following vectors span  $\mathbb{C}^M$ :

$$\{Ux\}_{U \in G} = \left\{\frac{M}{M+1}\Psi P\Psi^* x\right\}_{P \in S_{M+1}} = \{\Phi P\Phi^* x\}_{P \in S_{M+1}},$$

which is equivalent to having that  $\{Py\}_{P \in S_{M+1}}$  spans  $1_{M+1}^\perp$  for any nonzero  $y \in 1_{M+1}^\perp$ . This in turn is equivalent to showing that  $z = 0$  is the only choice of  $z \in 1_{M+1}^\perp$  for which  $\langle z, Py \rangle = 0$  for all permutations  $P$ . To do this, fix any indices  $n_1 \neq n_2, n_3 \neq n_4$  from  $\{1, \dots, M+1\}$ , and consider the zero inner product  $\langle z, P_1 y \rangle$  that arises from any permutation  $P_1$  which takes  $n_3$  to  $n_1$  and  $n_4$  to  $n_2$ . From this, now subtract the zero inner product from a permutation  $P_2$  which is identical to  $P_1$ , except that it takes  $n_3$  to  $n_2$  and  $n_4$  to  $n_1$ :

$$\begin{aligned} 0 &= \langle z, P_1 y \rangle - \langle z, P_2 y \rangle \\ &= z[n_1]\overline{y[n_3]} + z[n_2]\overline{y[n_4]} - z[n_1]\overline{y[n_4]} - z[n_2]\overline{y[n_3]} \\ &= (z[n_1] - z[n_2])(\overline{y[n_3]} - \overline{y[n_4]}). \end{aligned} \quad (13)$$

Now, since  $0 \neq y \in 1_{M+1}^\perp$  we have that  $y$  is a nonzero vector whose entries sum to zero, and so in particular there exists indices  $n_3$  and  $n_4$  such that  $y[n_3] - y[n_4] \neq 0$ . As such, (13) implies that  $z[n_1] = z[n_2]$  for every choice of  $n_1 \neq n_2$ , namely that the entries of  $z$  are all equal. Since  $z \in 1_{M+1}^\perp$ , this means  $z = 0$  as claimed. We summarize these facts below:

**Theorem 8.** *Let  $\Psi$  be an  $M \times (M+1)$  matrix whose unit columns form a regular simplex in  $\mathbb{C}^M$ . Let  $f = \sqrt{M/(M+1)}\Psi g$ , where  $g$  is any unit-norm vector  $g \in \mathbb{C}^{M+1}$  whose entries sum to zero. Then the group frame*

$$\{Uf\}_{U \in G} := \left\{ \frac{M}{M+1} \Psi P \Psi^* f \right\}_{P \in S_{M+1}}$$

*is a unit norm tight frame of  $(M+1)!$  elements for  $\mathbb{C}^M$ . Moreover, each row of the Gram matrix of this frame has entries of the form  $\{\langle f, Uf \rangle\}_{U \in G} = \{\langle g, Pg \rangle\}_{P \in S_{M+1}}$ . Here,  $P$  ranges over all  $(M+1) \times (M+1)$  permutation matrices.*

We now use these ideas to construct a frame to be used in conjunction with the bound (10), where  $d_F[k]$  denotes the  $k$ th largest value of the form  $|\langle f, Uf \rangle|^2 = |\langle g, Pg \rangle|^2$ . In particular, our goal is to find a unit norm vector  $g \in 1_{M+1}^\perp$  for which the average of the  $K$  largest values of  $d_F[k]$  is very close to the average of all of its values:  $\frac{1}{M}$ .

Moreover, considering the underlying application of NERFs, we prefer not to transmit as many as  $(M+1)!$  frame coefficients to convey an  $M$ -dimensional signal. For this reason, we seek vectors  $g$  which are fixed by a large subgroup of permutation matrices, namely, vectors with large level sets; this way, we can get away with only using representatives of distinct cosets of this large subgroup. In this paper, we only consider vectors of two level sets, say

$$g = (\underbrace{a, a, \dots, a}_{L \text{ times}}, \underbrace{b, b, \dots, b}_{M+1-L \text{ times}}). \quad (14)$$

Choosing  $g$  in this way guarantees that the corresponding group frame only has  $\binom{M+1}{L}$  distinct elements. Moreover, since each of these unique elements appears the same number of times, namely  $L!(M+1-L)!$  times, the  $\binom{M+1}{L}$ -element subframe is still tight.

To estimate the NERF properties of such frames using (10), we first need to find explicit expressions for  $a$  and  $b$ . Here, the condition  $\langle g, 1_{M+1} \rangle = 0$  implies  $La + (M+1-L)b = 0$ . Combining this with the fact that  $g$  has unit norm then gives

$$a = \sqrt{\frac{M+1-L}{(M+1)L}}, \quad b = -\sqrt{\frac{L}{(M+1)(M+1-L)}}, \quad (15)$$

where we take  $a > 0$  without loss of generality. Next, note that  $\langle g, Pg \rangle$  is completely determined by the number  $J$  of indices  $n$  for which  $g[n] = (Pg)[n] = a$ . This leads to the following calculation:

$$\langle f, Uf \rangle = \langle g, Pg \rangle = Ja^2 + 2(L-J)ab + (M+1+J-2L)b^2 = \frac{J(M+1) - L^2}{L(M+1-L)}.$$

Moreover, of the  $\binom{M+1}{L}$  distinct  $Uf$ 's in this construction, there are  $\binom{L}{J} \binom{M+1-L}{L-J}$  which produce the above inner product, since  $J$  of the  $a$ 's in  $Pg$  must align with  $a$ 's in  $g$ , while the other  $L-J$   $a$ 's in  $Pg$  align with  $b$ 's in  $g$ . In the special case where  $g$  has  $L = 2$   $a$ 's, we have a total of  $\binom{M+1}{2}$  distinct  $Uf$ 's, and the distribution of inner products is given by

$$\{\langle f, Uf \rangle\} = \begin{cases} 1 & \text{with multiplicity } 1, \\ \frac{M-3}{2(M-1)} & \text{with multiplicity } 2(M-1), \\ -\frac{2}{M-1} & \text{with multiplicity } \frac{1}{2}(M-1)(M-2). \end{cases} \quad (16)$$

As verified below, substituting this fact into (10) yields the following result:

**Theorem 9.** *Pick  $M \geq 7$  and consider the  $M \times \binom{M+1}{2}$  frame  $F$  with columns of the form  $\sqrt{M/(M+1)}\Psi Pg$ , where  $\Psi$  is an  $M \times (M+1)$  regular simplex and the  $Pg$ 's are distinct permutations of  $g$ , which is defined by (14) and (15) with  $L = 2$ . Then  $F$  is a  $(p, C)$ -numerically erasure-robust frame for every  $p \leq \frac{(C^2-1)^2}{(C^2+1)^2(M+1)}$ .*

The above guarantee bears a striking resemblance to Theorem 6, despite the distribution  $d_F$  being significantly different. Again, while this result is not nearly as good as the ones we got for ETFs or random frames, it still gives something; for example, removing any  $0.48M$  of the  $\binom{M+1}{2}$  frame vectors will leave a submatrix of condition number smaller than 10. As one would expect, there are similar NERF results for the frames that correspond to larger values of  $L$ , but we do not report them here.

*Proof of Theorem 9.* Since  $M \geq 7$ , the sizes of the inner products in (16) are nonincreasing, and so  $d_F$  is defined accordingly. Also, taking  $K = (1-p)N$  with  $p \leq \frac{(C^2-1)^2}{(C^2+1)^2(M+1)} \leq \frac{1}{M+1}$ , we claim that  $K \geq 2(M-1) + 1$ . Indeed,

$$K \geq \left(1 - \frac{1}{M+1}\right)N = \frac{M^2}{2} \geq 2(M-1) + 1,$$

where the last inequality follows from  $M \geq 7 \geq 2 + \sqrt{2}$ . Since  $K \geq 2(M-1) + 1$ , then applying (10) to (16) yields

$$\begin{aligned} \delta_K^2 &\leq \frac{M^2}{K} \left( \sum_{k=1}^K d_F[k] - \frac{K}{M} \right) \\ &= \frac{M^2}{K} \left( 1 + 2(M-1) \left( \frac{M-3}{2(M-1)} \right)^2 + (K - (2M-1)) \left( \frac{2}{M-1} \right)^2 - \frac{K}{M} \right) \\ &= \left( \frac{M(M+1) - 2K}{2K} \right) \left( \frac{M(M^2 - 6M + 1)}{(M-1)^2} \right). \end{aligned}$$

Since  $K = (1-p)N$  and  $N = \binom{M+1}{2}$ , we can simplify to get

$$\delta_K^2 \leq \frac{pM(M^2 - 6M + 1)}{(1-p)(M-1)^2} \leq \frac{pM}{1-p}.$$

From here,  $p \leq \frac{(C^2-1)^2}{(C^2+1)^2(M+1)}$  and (11) together imply  $\delta_K^2 \leq \frac{(C^2-1)^2}{(C^2+1)^2}$ , which we substitute into (3) to conclude that  $\text{Cond}(F_K) \leq C$ .  $\square$

### 3. LIMITING OUR EXPECTATIONS

The previous section gave four different constructions of numerically erasure-robust frames. The last three constructions were deterministic, and their proofs hinged on how coherent a subcollection of frame vectors can be. In this section, we shed some light on the fundamental limits of NERFs by again considering the coherence of frame subcollections. We start with the following lemma, which says that a matrix with similar columns will have a large condition number:

**Lemma 10.** *Take an  $M \times N$  matrix  $F$  with unit-norm columns. Then for every unit vector  $x \in \mathbb{R}^M$ ,*

$$(\text{Cond}(F))^2 \geq \frac{(M-1)\|F^*x\|^2}{N - \|F^*x\|^2}.$$

*Proof.* First, we have  $\lambda_{\max}(FF^*) = \|F^*\|_2^2 \geq \|F^*x\|^2$ . Next, take  $\{x_m\}_{m=1}^M$  to be some orthonormal basis with  $x_1 = x$ . Then  $\lambda_{\min}(FF^*) \leq \|F^*x_m\|^2$  for every  $m$ , and so averaging over  $m = 2, \dots, M$  gives

$$\lambda_{\min}(FF^*) \leq \frac{1}{M-1} \sum_{m=2}^M \|F^*x_m\|^2 = \frac{1}{M-1} \sum_{n=1}^N \sum_{m=2}^M |\langle x_m, f_n \rangle|^2.$$

Since each  $f_n$  has unit norm and  $\{x_m\}_{m=1}^M$  is an orthonormal basis with  $x_1 = x$ , we continue:

$$\lambda_{\min}(FF^*) \leq \frac{1}{M-1} \sum_{n=1}^N \left(1 - |\langle x, f_n \rangle|^2\right) = \frac{N - \|F^*x\|^2}{M-1}.$$

Combining this with our lower bound on  $\lambda_{\max}(FF^*)$  gives the result.  $\square$

To be explicit, the lower bound in Lemma 10 is exceedingly large when the columns of  $F$  each have a large inner product with  $x$ . We now use this lemma to prove the following statement on the fundamental limits of NERFs:

**Theorem 11.** *Take a sequence of real  $M \times N_M$  frames  $\{F_M\}_{M=1}^\infty$ , pick  $C > 1$ , and take a sequence of erasure rates  $\{p_M\}_{M=1}^\infty$  such that*

$$\liminf_{M \rightarrow \infty} p_M > 1 - 2Q(C), \quad Q(t) := \frac{1}{\sqrt{2\pi}} \int_t^\infty e^{-u^2/2} du. \quad (17)$$

*Then for all sufficiently large  $M$ ,  $F_M$  is not a  $(p_M, C)$ -numerically erasure-robust frame.*

*Proof.* For notational simplicity, we write  $F = F_M$ ,  $N = N_M$  and  $p = p_M$ . Further let  $\mathbb{S}^{M-1}$  denote the unit sphere in  $\mathbb{R}^M$ . For any  $x \in \mathbb{S}^{M-1}$ , consider the “polar caps” of the sphere about  $\pm x$ , namely the set  $B(x) := \{y \in \mathbb{S}^{M-1} : |\langle x, y \rangle|^2 \geq \frac{C^2}{M}\}$ . For any such *bi-cap*, we may count the number of frame elements that it contains, namely the cardinality of the set  $B(x) \cap \{f_n\}_{n=1}^N$ . Let  $x_0$  denote the point on the sphere whose bi-cap contains the most frame elements. By the pigeonhole principle, the fraction of frame elements contained in this bi-cap is at least the fraction of its surface area to the surface area of the entire sphere:

$$\left| B(x_0) \cap \{f_n\}_{n=1}^N \right| \geq N \cdot \frac{\text{Area}(B(x))}{\text{Area}(\mathbb{S}^{M-1})}.$$

Assuming for the moment that  $1 - p \leq \text{Area}(B(x))/\text{Area}(\mathbb{S}^{M-1})$ , we may take  $\mathcal{K}$  to be the indices of any  $K = (1 - p)N$  of the  $f_n$ ’s in  $B(x_0) \cap \{f_n\}_{n=1}^N$ . Then

$$\|F_{\mathcal{K}}^*x_0\|^2 = \sum_{k \in \mathcal{K}} |\langle x_0, f_k \rangle|^2 \geq K \frac{C^2}{M},$$

and so applying Lemma 10 to the  $M \times K$  matrix  $F_K$  gives

$$(\text{Cond}(F_K))^2 \geq \frac{(M-1)\|F_K^* x_0\|^2}{K - \|F_K^* x_0\|^2} \geq \frac{(M-1)K \frac{C^2}{M}}{K - K \frac{C^2}{M}} = \frac{M-1}{M-C^2} C^2 > C^2,$$

as claimed. Thus, it only remains to show that

$$1 - p \leq \frac{\text{Area}(B(x))}{\text{Area}(\mathbb{S}^{M-1})} \quad (18)$$

for sufficiently large  $M$ .

To this end, pick  $M$  large enough so that  $\frac{C^2}{M} < 1$  and take  $\theta \in (0, \frac{\pi}{2})$  such that  $\cos^2 \theta = \frac{C^2}{M}$ . Then  $B(x)$  is the union of both polar caps of angular radius  $\theta$  centered at  $\pm x$ . Using hyperspherical coordinates, we find that

$$\text{Area}(B(x)) = 2 \text{Area}(\mathbb{S}^{M-2}) \int_0^\theta \sin^{M-2} \varphi \, d\varphi. \quad (19)$$

Next, we can substitute  $t = \cos \varphi$  to get

$$\int_0^\theta \sin^{M-2} \varphi \, d\varphi = \int_0^\theta \sin^{M-3} \varphi \sin \varphi \, d\varphi = \int_{\cos \theta}^1 (1-t^2)^{\frac{M-3}{2}} dt. \quad (20)$$

Note that the area of  $\mathbb{S}^{M-1}$  is given by replacing  $\theta$  with  $\frac{\pi}{2}$  in (19) and (20), and so

$$\frac{\text{Area}(B(x))}{\text{Area}(\mathbb{S}^{M-1})} = \frac{\int_{\cos \theta}^1 (1-t^2)^{\frac{M-3}{2}} dt}{\int_0^1 (1-t^2)^{\frac{M-3}{2}} dt}.$$

Substituting  $u = t\sqrt{M-3}$  and recalling that  $\cos^2 \theta = \frac{C^2}{M}$  results in new integrals which converge as  $M$  grows large:

$$\frac{\text{Area}(B(x))}{\text{Area}(\mathbb{S}^{M-1})} = \frac{\int_{C\sqrt{\frac{M-3}{M}}}^{\sqrt{M-3}} \left(1 - \frac{2}{M-3} \frac{u^2}{2}\right)^{\frac{M-3}{2}} du}{\int_0^{\sqrt{M-3}} \left(1 - \frac{2}{M-3} \frac{u^2}{2}\right)^{\frac{M-3}{2}} du}.$$

Specifically, since  $(1 + \frac{x}{n})^n$  converges from below to  $e^x$  for all  $x \geq 0$ , we can apply the Lebesgue dominated convergence theorem to the Gaussian to obtain

$$\frac{\text{Area}(B(x))}{\text{Area}(\mathbb{S}^{M-1})} \longrightarrow \frac{\int_C^\infty e^{-u^2/2} du}{\int_0^\infty e^{-u^2/2} du} = 2Q(C).$$

This implies that as  $M$  grows large, our assumption (17) guarantees (18), as needed.  $\square$

As a corollary to Theorem 11, note that if  $p_M \rightarrow 1$  as  $M$  gets large, then the worst-case condition number diverges to infinity. Specifically, this establishes that  $M \times N$  full spark frames with  $M = o(N)$  cannot be “maximally robust to erasures” in a numerical sense; for sufficiently large  $M$ , the adversary can delete  $N - M$  columns of the frame in a way that leaves an arbitrarily ill-conditioned square submatrix. This highlights the value of a theory of numerically erasure-robust frames.

## 4. IMPLICATIONS AND REMAINING PROBLEMS

Having constructed several numerically erasure-robust frames, and having further proved certain fundamental limits, we conclude with a few interesting observations. First, we consider an implication for maximal ETFs: no  $M \times N$   $(p, C)$ -NERF can have  $(1 - p)N$  zeros in a common row, since otherwise the adversary can delete the other  $pN$  columns and leave a rank-deficient submatrix. Since Theorem 4 also applies to maximal ETFs, this implies that there is no basis over which half of a maximal ETF's vectors share a common zero coordinate. That is, if maximal ETFs exist, then they cannot be too sparse in any basis.

Due to their computational benefits, frames which have a sparse representation have recently become a subject of active research [8, 10]. In this vein, one attractive feature of Steiner ETFs is their naturally sparse representation; in fact, the proportion of nonzero entries in an  $M \times N$  Steiner ETF is  $O(M^{-1/2})$  [15]. However, no Steiner ETF can be maximal, for they have at most  $N = O(M^{3/2})$ . The work presented here reinforces this fact: since no  $M \times N$   $(p, C)$ -NERF can be very sparse, and since ETFs with  $N = \Omega(M^2)$  are NERFs by Theorem 5, we see that neither Steiner ETFs—nor any generalization of the Steiner construction with similar levels of sparsity—will ever be able to produce ETFs in which  $N = \Omega(M^2)$ .

Recall that  $M \times N$  full spark frames have the defining property that every subcollection of  $M$  columns spans; trivially, this implies that every subcollection of size *at least*  $M$  also spans. By analogy, it is natural to ask whether a  $(p, C)$ -NERF is also a  $(p', C)$ -NERF for every  $p' \in [0, p)$ . However, it is not clear whether this is the case, since deleting columns does not necessarily worsen a frame's conditioning. As an example, the union of an orthonormal basis with some unit vector is not as well conditioned as the orthonormal basis which survives the deletion of the last vector. While this open question is interesting, it is inconsequential in practice: If the adversary deletes less than  $pN$  of the frame vectors, we can neglect more of them to guarantee a well-conditioned subframe.

Another remark: Reviewing the results of this paper, we know there exist NERFs with  $p < \frac{1}{2}$  by Theorem 4. Meanwhile, Theorem 11 states that for any fixed  $C$ , there do not exist NERFs with values of  $p$  that grow arbitrarily close to 1. Various questions remain: Do there exist NERFs with  $p \in [\frac{1}{2}, 1)$ ? If so, what is the largest  $p$  for which  $(p, C)$ -NERFs exist? Interestingly, this “one-half barrier” appears to be more than a mere artifact of Theorem 4. To be clear, every matrix  $F$  whose entries are  $\pm 1$ 's cannot be a NERF with  $p \geq \frac{1}{2}$ ; for any two rows of  $F$ , the corresponding entries are either equal or opposite, and so the adversary can delete the columns corresponding to the less popular relationship and leave a rank-deficient matrix. Moreover, random matrix methods [6, 21] apply to matrices of  $\pm 1$  entries without loss of effectiveness, and so breaking the one-half barrier, if it is even possible, will likely require other methods.

## REFERENCES

- [1] B. Alexeev, J. Cahill, D.G. Mixon, Full spark frames, submitted, Available online: arXiv:1110.3548
- [2] W.O. Alltop, Complex sequences with low periodic correlations, IEEE Trans. Inform. Theory, 26 (1980) 350–354.
- [3] D.M. Appleby, Symmetric informationally complete-positive operator valued measures and the extended Clifford group, J. Math. Phys. 46 (2005) 052107/1–29.

- [4] B. Balan, B.G. Bodmann, P.G. Casazza and D. Edidin, Painless Reconstruction from Magnitudes of Frame Vectors, *J. Fourier Anal. Appl.* 15 (2009) 488–501.
- [5] A.S. Bandeira, M. Fickus, D.G. Mixon, P. Wong, The road to deterministic matrices with the restricted isometry property, submitted, Available online: [arXiv:1202.1234](https://arxiv.org/abs/1202.1234)
- [6] R. Baraniuk, M. Davenport, R. DeVore, M. Wakin, A simple proof of the restricted isometry property for random matrices, *Constr. Approx.* 28 (2008) 253–263.
- [7] J.J. Benedetto, M. Fickus, Finite normalized tight frames, *Adv. Comput. Math.* 18 (2003) 357–385.
- [8] R. Calderbank, P.G. Casazza, A. Heinecke, G. Kutyniok, A. Pezeshki, Sparse fusion frames: Existence and construction, *Adv. Comput. Math.* 35 (2011) 1–31.
- [9] P.G. Casazza, M. Fickus, Fourier transforms of finite chirps, *EURASIP J. Appl. Signal Process.* 2006 (2006) 70204/1–7.
- [10] P.G. Casazza, A. Heinecke, F. Krahmer, G. Kutyniok, Optimally sparse frames, *IEEE Trans. Inform. Theory* 57 (2011) 7279–7287.
- [11] P.G. Casazza, J. Kovačević, Equal-norm tight frames with erasures, *Adv. Comput. Math.* 18 (2003) 387–430.
- [12] E.J. Candès, The restricted isometry property and its implications for compressed sensing, *C. R. Acad. Sci. Paris, Ser. I* 346 (2008) 589–592.
- [13] E.J. Candès and T. Tao, Decoding by linear programming, *IEEE Trans. Inform. Theory* 44 (2005) 4203–4215.
- [14] K.R. Davidson and S.J. Szarek, Local operator theory, random matrices and Banach spaces, In: *Handbook in Banach Spaces Vol I*, ed. W.B. Johnson, J. Lindenstrauss, Elsevier (2001), 317–366.
- [15] M. Fickus, D.G. Mixon and J.C. Tremain, Steiner equiangular tight frames, *Linear Algebra Appl.* 436 (2012) 1014–1027.
- [16] V.K. Goyal, J. Kovačević, J.A. Kelner, Quantized frame expansions with erasures, *Appl. Comp. Harmon. Anal.* 10 (2001) 203–233.
- [17] R.B. Holmes, V.I. Paulsen, Optimal frames for erasures, *Linear Algebra Appl.* 377 (2004) 31–51.
- [18] D. Jungnickel, A. Pott, K.W. Smith, Difference sets, In: C.J. Colbourn, J.H. Dinitz (Eds.), *Handbook of Combinatorial Designs* (2007) 419–435.
- [19] M. Planat, H.C. Rosu, S. Perrine, A survey of finite algebraic geometrical structures underlying mutually unbiased quantum measurements, *Found. Phys.* 36 (2006) 1662–1680.
- [20] M. Püschel, J. Kovačević, Real, tight frames with maximal robustness to erasures, *Proc. Data Compr. Conf.* (2005) 63–72.
- [21] M. Rudelson, R. Vershynin, On sparse reconstruction from Fourier and Gaussian measurements, *Comm. Pure Appl. Math.* 61 (2008) 1025–1045.
- [22] T. Strohmer and R.W. Heath, Grassmannian frames with applications to coding and communication, *Appl. Comp. Harmon. Anal.* 14 (2003) 257–275.
- [23] R. Vale, S. Waldron, Tight frames and their symmetries, *Constr. Approx.* 21 (2005) 83–112.
- [24] S. Waldron, On the construction of equiangular frames from graphs, *Linear Algebra Appl.* 431 (2009) 2228–2242.
- [25] L.R. Welch, Lower bounds on the maximum cross correlation of signals, *IEEE Trans. Inform. Theory* 20 (1974) 397–399.
- [26] P. Xia, S. Zhou and G.B. Giannakis, Achieving the Welch bound with difference sets, *IEEE Trans. Inform. Theory* 51 (2005) 1900–1907.

(Fickus) DEPARTMENT OF MATHEMATICS AND STATISTICS, AIR FORCE INSTITUTE OF TECHNOLOGY, WRIGHT-PATTERSON AIR FORCE BASE, OH 45433, USA; [MATTHEW.FICKUS@AFIT.EDU](mailto:MATTHEW.FICKUS@AFIT.EDU)

(Mixon) PROGRAM IN APPLIED AND COMPUTATIONAL MATHEMATICS, PRINCETON UNIVERSITY, PRINCETON, NEW JERSEY 08544, USA; E-MAIL: [DMIXON@PRINCETON.EDU](mailto:DMIXON@PRINCETON.EDU)